

Detection of Common Errors in Turkish EFL Students' Writing through a Corpus Analytic Approach

Elif Tokdemir Demirel¹

¹ Department of Applied English and Translation Studies, Ufuk University, Ankara, Turkey

Correspondence: Elif Tokdemir Demirel, Department of Applied English and Translation Studies, Ufuk University, 06836, İncek, Ankara, Turkey. Tel: 90-312-586-7498. E-mail: elif.demirel@ufuk.edu.tr

Received: July 22, 2017 Accepted: September 18, 2017 Online Published: September 22, 2017

doi: 10.5539/elt.v10n10p159 URL: <http://doi.org/10.5539/elt.v10n10p159>

Abstract

The present study aims to explore Turkish EFL students' major writing difficulties by analyzing the frequent writing errors in academic essays. Accordingly, the study examined errors in a corpus of 150 academic essays written by Turkish EFL students studying at the Department of English Language and Literature at a public university in Turkey. The essays were written on assigned topics as take home exam papers or assignments in the context of a first year academic writing course. The corpus consisted of essays of various lengths ranging from 500 word essays to 1500 word essays. The essays were compiled into a corpus and analyzed by using a concordance program. The essays were also checked for plagiarism using the online plagiarism detection software and plagiarized essays were excluded from the analysis. Errors were classified by using an error classification system which was organized according to lexico-grammatical categories. The resulting categories consisted of mostly syntactic and lexical categories of error but academic style errors were considered as well. As a result of the analysis, in terms of error categories, the most frequent errors were observed in the verb related error categories. When considered individually, the most frequent errors were observed in noun modification and were mostly interference related.

Keywords: academic writing; corpus linguistics; error analysis; intralingual error, interference error

1. Introduction

Non-native speakers of English inevitably make errors in writing and usually knowledge of grammatical errors does not guarantee the production of error-free language (Bowden & Fox, 2002). Especially in productive skills such as writing and speaking, it is difficult for non-native learners to produce accurate and fluent language. Turkish non-native learners of English are not an exception. Turkish learners of English have difficulties in productive skills, especially writing since writing in English is not a skill which is emphasized prior to tertiary level education.

In most high schools in Turkey, the major skills which are exercised are grammar and reading. However, at university level, especially at departments where English is the main area of study such as English Literature Departments or English Language Teaching Departments, it becomes inevitable for learners to have a good command of English for all four skills: reading, writing, speaking and listening. In case of the present research, the context of research is the English Language and Literature Department at Karadeniz Technical University. As students studying English Literature, students at this department are expected to express themselves in writing accurately and fluently and using an academic language. A graduation requirement is to write an extended research in the form of a graduation thesis which has high expectations from students in terms of using academic language appropriately. For this reason, the most significant rationale behind this study is to explore the difficulties of learners in using academic English in their writing to provide them with effective guidance. By putting together a tagged error corpus the researcher aims at building a rich source of materials development which will assist academic writing classes at the department.

Accordingly, Gaskell and Cobb (2004) argue that using a corpus based approach has several benefits in aiding students to overcome their writing errors such as by increasing the number of examples that L2 learners are exposed to in a given unit of time, by organizing examples so their patterns are highlighted, by getting learners to attend to the examples, and by providing systematic feedback on the success of interpreting the examples. A properly configured concordance, with the help of an error corpus, thus can be very useful in helping learners

with overcoming their difficulties with academic writing.

Despite the growing interest in computer aided error analysis of non-native students' writing in the world, there are limited number of studies in Turkey which analyze Turkish students writing with a corpus-based approach (Kırkgöz, 2010; Erkaya, 2012; Özhan, 2012). Therefore, in order to fill this gap, this study aims to explore 1st year Turkish university students' major writing problems by analyzing the nature and distribution of their writing errors by using CEA approach. The second purpose is to discuss the possible sources of errors according to Richards' (1974) classification of sources of error. Richards (1974) differentiates three sources of error. The first source of error is called "interference error" or "interlingual error", which results from the mother tongue interference. The second source, "intra-lingual error", reflects the incorrect generalization of the rules within the target language. The last source is "developmental errors", occurring when the learners hypothesize about the target language based on their limited knowledge.

In this context, the study does not aim at solely error correction, but it aims at a clear description of learner performance to guide teaching practices and to provide data about non-native writing for other researchers. Researchers have opposing views on the effects of corrective feedback on the performance of student writers. While Truscott (2007) supports the view that correction does not have an effective role on the reduction of errors in writing and even harms the learning process, Ferris (2004) recommends the use of error correction in writing. The rationale behind researching errors in student writing has been changing over the years. For example, in the early years of writing research, error correction was seen as a way of reaching the 'ideal and flawless' performance in writing; however today detection of errors are used as a first step towards understanding the process of writing, identifying the learners and creating computerized aids which could help students during the process in correcting and revising their writing. For this reason, it is important to detect problematic areas of specific groups of learners from different native language backgrounds before we can offer a remedy for better performance in writing to non-native learners of English.

EA (Error Analysis) approach which was at its heyday in the 1970s opened a window onto learners' interlanguage and produced a wealth of error typologies. However, it has been reported to have certain limitations such as:

- Heterogeneous learner data
- Fuzzy categories
- Inability to cater for phenomena such as avoidance
- Being restricted to what the learner cannot do
- Giving a static picture of L2 learning (Dagneaux, Denness, & Granger, 1998)

With the emergence of CLC (Computer Learner Corpora) in the 1990s, a new source of data was presented for SLA research. Identification of common learner difficulties through CLC has led to the development of pedagogic materials focusing on specific difficulties of learners. For example, dictionaries which make use of learner corpora started to incorporate results from corpus analysis (Nesselhauf, 2004).

As an alternative to the traditional EA, CEA (Computer Aided Error Analysis) provides certain advantages. As a contribution to Contrastive Analysis, CEA suggests that comparing learners L1 with the target language to determine areas of difficulty is not sufficient but that the best way is to find out what these difficulties are through the analysis of learner language and then compare it with native speaker production (Nesselhauf, 2004).

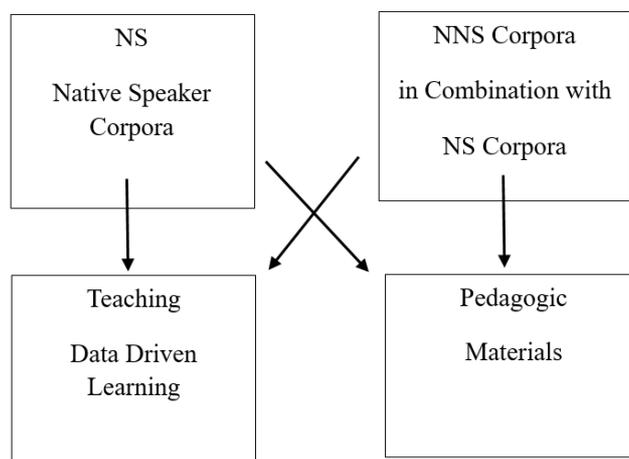


Figure 1. The CEA approach (adapted from Nesselhauf, 2004)

2. Literature Review

Researchers have been using CEA method and learner corpora to develop error detection systems which could work with non-native language since, grammar checkers are usually built to understand language by native speakers (Granger & Meunier, 1994; Milton, 1994; Dagneaux, Denness, & Granger, 1998). For example, Bowden and Fox (2002) have developed the GRADES, an expert system to identify and diagnose common syntactic errors among non-native speakers, particularly Japanese-speaking learners of English. Another widely used error classification system is the Error tagging Manual, version 1.1. (Dagneaux, Denness, Granger, & Meunier, 1996). Both classification systems have been utilized in the present study to develop a unique classification system which could be used for analyzing errors of Turkish EFL learners in academic writing.

Many researchers have studied errors in non-native students' writing from different first language backgrounds: Spanish, Chinese, Romanian, Iranian, Algerian and Ghanaian (Dagneaux, Denness, & Granger, 1998; Huaqing, 2016; Punga & Parlog, 2015; Nezami & Najafi, 2012; Mammeri, 2016; Adjedi, 2015) to name a few and with different research purposes such as exploring errors in the use of grammatical (Han, Chodorow, & Leacock, 2006) or lexical items (Altenberg & Granger, 2001; Granger & Tyson, 1996).

The findings of these studies are briefly summarized in this section. A number of the studies which utilize an error-tagged corpus have served the purpose of defining a learner population. Such a study is the pioneering study with an error tagged corpus by Dagneaux, Denness and Granger (1998). The researchers used a fully error-tagged French learner corpus (150,000 words) and were able to characterize the learner population in terms of the proportion of the major error categories. They found three main areas of grammatical difficulty: articles, verbs and pronouns. Each of these categories accounted for approximately a quarter of the grammatical errors (27% for articles and 24% for pronouns and verbs, respectively). Doolan and Miller's (2012) study also aims to characterize a specific generation of students in the US educational system: the so-called generation 1.5 students who "(a) have been in the US educational system for more than four years, (b) regularly speak a language other than English at home, (c) have relatively strong English speaking and listening skills, (d) are younger than 25 years old." (p. 1).

Other studies on the subject of error analysis aimed at finding the most problematic areas through analysis of errors. For example, Diez-Bedmar (2011) used CEA to analyze the errors of Spanish students when writing in English for the University Entrance Exam. The study focused on eight error categories: Form, Grammar, Lexis, Punctuation, Register, Style, Word and Lexico-Grammar. The highest mean of errors was in Grammar, followed by Lexis and Form. The first four types of most common errors found were the proper selection of vocabulary, spelling errors, use of pronouns and articles.

Rather than researching a wide variety of errors, some studies focused on specific language items which were diagnosed to be potential sources of error or variation between L1 and L2 writers. Flowerdew (2010) compared the use of signaling nouns across L1 and L2 learner corpora using the ICLE Locness (L1 writers) corpus as a reference corpus and a learner English corpus written by Cantonese speaking learners of English. Paquot (2013) investigated the effect of L1 transfer on French EFL learners' use of lexical bundles in writing. Huaqing's (2016)

study with Chinese Learners of English employed CA and CEA methods and focused on word form errors specifically. The word form errors were found to account for 28.42% of the total language errors. Punga and Parlog (2015) analyzed errors resulting from the interference between the learners' mother tongue (Romanian) and English as a foreign language using a corpus based methodology. They classified errors according to the type of L1-L2 transfer involved in their production. Among the most frequent errors found word order errors accounted for 26.21 percent of the sample, misuse of articles accounted for 25.24 percent and vocabulary errors accounted for 19.42 percent and errors in the use of prepositions accounted for 16.50 percent. Xia (2015) analyzed Chinese college students' writing in terms of word class errors. In the study, the word class errors were found to be more frequent than errors in collocations and were very common. Mammeri (2015) analyzed written compositions of Algerian EFL students at the level of morpho-syntax using a corpus of 120 written compositions. The morpho-syntactic errors found were related to word order, subject-verb agreement, verb structure, noun/adjective/adverb structure, word/morpheme addition, word/morpheme omission, short forms/abbreviations, and conversational informal words respectively. In addition to detecting error patterns and possible sources of error, error tagged corpora have also been used to examine L2 developmental patterns (Thewissen, 2013). Thewissen's study used 40 error types in order to trace the type of development along the B1–C2 proficiency range of the Common European Framework.

The influence of L1 interference is seen as one of the major causes of errors in second language writing. A number of studies have concluded that most errors observed in learner written production are caused by L1 interference (Chuang & Nessi, 2006; Diez-Bedmar & Papp, 2008; Hawkins & Buttery, 2010). This evidence has even led researchers to start using CEA as a tool for automatic detection of learners' L1 (Bestgen, Granger, & Thewissen, 2012; Crossley & McNamara, 2012).

3. Methodology

This part mainly deals with the methodology used in this study, which consists of 5 sections, namely research design, data collection procedure, participants, research questions and research instruments.

3.1 Research Design

The present study is descriptive in nature and uses a corpus-based methodology. As Kennedy (1998) reports: "a major reason for collecting linguistic corpora is to provide the basis for more accurate and reliable descriptions of how languages are structured and learned" (p. 88). The corpus-based description in this study comprises of two levels; at one level the distribution of linguistics features is provided through tagging of the corpus, at the other level frequencies of errors in the use of certain lexico-grammatical elements of language are provided.

3.2 Data Collection Procedure

The study made use of a corpus of student essays. The corpus used in the study is composed of 150 academic essays (99,352 words) of different types, extended argument, argument, definition and process essays written by 1st year students at Karadeniz Technical University, Department of English Language and Literature in the scope of an academic writing course. For the selection of the essays, convenience sampling was used and all essays except for ones in which plagiarism was detected were included in the study in order to have a large enough corpus of student writing.

Table 1 shows the distribution of essay types included in the corpus. Among the essays used to compile a corpus, 30 were extended argument essays, 43 were argumentative essays, 42 were process essays and 35 were definition essays. The subjects of the essays were chosen by the students for argumentative essays, process essays and definition essays. The extended argument essay was written by the students as a take-home exam and required students to cite relevant literature in support of their view. The length of this essay ranged from 3 to 5 double spaced pages excluding references. The argumentative essay was in the format of a five paragraph essay in which students argued their view with supporting ideas around a central thesis statement. The process essay was also a five-paragraph essay in which students described the steps in a process in detail. The definition essay required students to define a concept of their choice in a clear and concise manner in a five paragraph essay.

Some of the subjects chosen by the students are: "How to prepare for European Travel?" for a process essay, "The source of life: music" for a definition essay, "Does Facebook violate privacy?" for an argumentative essay. For the extended argument essay, the subject was assigned by the teacher: "Does education effect national development?".

Table 1. Description of the learner corpus used in the study

type of essay	number of essays	number of word types	number of word tokens
extended argument essay	30	3892	29981
argumentative essay	43	2417	19327
process essay	42	3061	20063
definition essay	35	2511	29981
Total	150	11881	99352

3.3 Participants

The participants of the study were 45 students of which 10 are male and 35 are female studying at the 1st year at KTU, department of English Language and Literature Department. The ages of the participants ranged from 19 to 21. All students in the first year were included in the study conveniently.

3.4 Research Questions

The study aimed at answering the following research questions:

- 1) What are the major lexico-grammatical and stylistic errors in the writing of non-native Turkish EFL students?
- 2) What is the nature of these lexico-grammatical error and stylistic errors: intralingual errors or interlingual errors?

3.4 Research Instruments

In this study, the GRADES (Bowden & Fox, 2002) system was used in combination with a corpus based approach to classify the writing errors of Turkish EFL students because this system is prepared based on non-native speakers' writing. The GRADES system seeks 17 error types classified in terms of lexico-grammar as such: Verb related errors (13 types) and Noun related errors (4 types). Essays were hand tagged using an adapted version of the GRADES error classification system. A new category was added for APA reference style and some parts of speech. Each error category was assigned an error tag. The first part of the error tag specified the general error category, and the second part specified the sub-category of error. For example, the tag 'VRE-SV' signifies, verb related error as the general error category and subject verb agreement as the sub-category.

The error classification system used in GRADES was adapted in this study to include more error types that were necessary for the tagging. All essays in the corpus were read and hand tagged for errors. During the error tagging process, it was realized that new categories were needed for the tagging to get a more detailed picture of the errors in the learner essays. For this reason, new categories were added for other lexico-grammatical categories such as pronouns, adjectives, prepositions. Other two categories related to style that were added were task specific since the learners were required to follow APA style guidelines when writing their essays, and since they wrote academic essays clausal type errors such as fragments and run-on sentences were also given importance. As a result, not only word-level errors, but also clausal errors and style errors were detected.

A second error classification system that was consulted was the Error Tagging Manual, version 1.1. (Dagneaux, Denness, Granger, & Meunier, 1996). This error taxonomy has been created based on frequent errors by students with a Romance language background. Seven main error categories made in this taxonomy are Form (F), Grammar (G), Lexico-Grammar (X), Lexis (L), Word Redundant, Word Missing and Word Order (W), Register (R) and Style. The taxonomy for the present study was developed by adapting these categories to the common errors of Turkish EFL students. The resulting error categories are presented in Appendix A. The two error classification systems: Grades (Bowden & Fox, 2002) and Error Tagging Manual, version 1.1 (Dagneaux, Denness, Granger, & Meunier, 1996) were combined in order to have an extensive classification which could cover all problem areas in student writing.

Before carrying out the error tagging, the learner corpus was tagged using the CLAWS (the Constituent Likelihood Automatic Word-tagging System) (Garside, 1987) online tagger (Appendix B) in order to get an overall idea about the language use of the learners in their essays. Figure 2 summarizes the results of the CLAWS tagging. This tagging system was preferred since it has been reported to have achieved 96-97% accuracy and is easily accessible for researchers through the internet. The CLAWS has two tagsets available online: UCREL CLAWS5 and UCREL CLAWS7. Because of the relatively small size of the learner corpus used

in the study, the smaller tagset was preferred for the tagging.

4. Results and Discussion

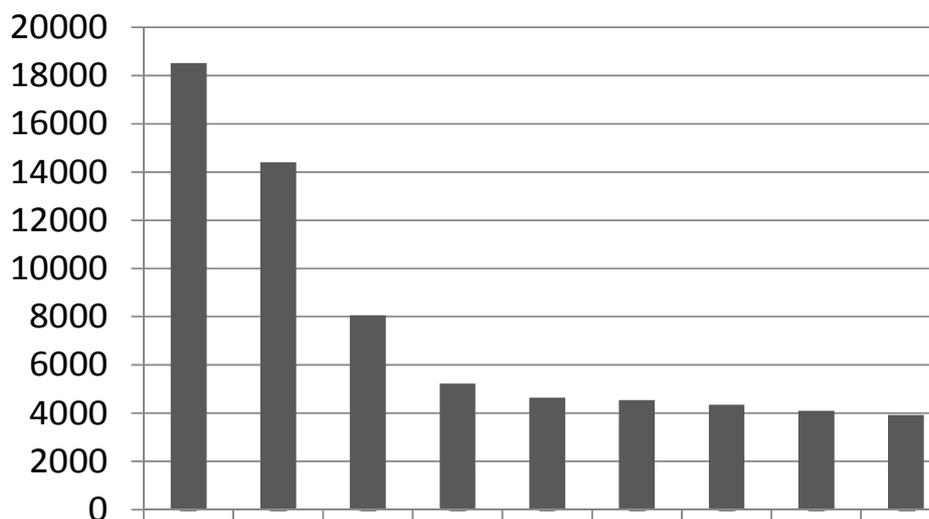


Figure 2. Results of the CLAWS tagging

The overall results of the tagging shows that the most frequently used categories in the learner corpus are nouns, verbs and prepositions, and the least frequent ones are determiners, adverbs and pronouns. Table 6 shows the types of nouns that were used in the learner corpus. Among all the nouns used in the corpus 67.2% is singular nouns, 21.8% is neutral nouns, 6.8% is plural nouns and 4.1% is proper nouns.

Table 2. Distribution of noun types in the learner corpus

Noun type	Code	Frequency in corpus	%
singular noun (e.g. PENCIL, GOOSE)	NN1	12454	67.2
plural noun (e.g. PENCILS, GEESE)	NN2	4045	21.8
noun (neutral for number) (e.g. AIRCRAFT, DATA)	NN0	1268	6.8
proper noun (e.g. LONDON, MICHAEL, MARS)	NP0	752	4.1
Total		18519	

The results of the tagging shows that a total of approximately 14.000 verbs are used in the learner corpus. Among these verbs, the most common three are infinitive of lexical verb (17.5%), base form of lexical verb (13.8%) and modal auxiliary verbs (12.4%), followed by -ing- forms of lexical verbs and -s forms of lexical verbs (11.5%), i.e. present simple verbs (8.6%). This distribution points to a rhetorical style which is argumentative and persuasive, since the purpose of the essays is to discuss an issue under discussion by using evidence from outside sources and writers' own interpretation.

Table 3. Distribution of verb types in the learner corpus

Verb Type	Code	Frequency in Corpus	%
infinitive of lexical verb	VVI	2519	17.5
base form of lexical verb (except the infinitive)(e.g. TAKE, LIVE)	VVB	1986	13.8
-s form of the verb "BE", i.e. IS, 'S	VBZ	1793	12.4
modal auxiliary verb (e.g. CAN, COULD, WILL, 'LL)	VM0	1650	11.5
-ing form of lexical verb (e.g. TAKING, LIVING)	VVG	1237	8.6
-s form of lexical verb (e.g. TAKES, LIVES)	VVZ	1062	7.4
past participle form of lex. verb (e.g. TAKEN, LIVED)	VVN	943	6.5
the "base forms" of the verb "BE" (except the infinitive), i.e. AM, ARE	VBB	727	5.0
infinitive of the verb "BE"	VBI	659	4.6
past tense form of lexical verb (e.g. TOOK, LIVED)	VVD	314	2.2
base form of the verb "HAVE" (except the infinitive), i.e. HAVE	VHB	296	2.1
base form of the verb "DO" (except the infinitive), i.e.	VDB	244	1.7
-s form of the verb "HAVE", i.e. HAS, 'S	VHZ	217	1.5
infinitive of the verb "HAVE"	VHI	128	0.9
past form of the verb "BE", i.e. WAS, WERE	VBD	123	0.9
-ing form of the verb "BE", i.e. BEING	VBG	114	0.8
infinitive of the verb "DO"	VDI	111	0.8
-s form of the verb "DO", i.e. DOES	VDZ	84	0.6
-ing form of the verb "DO", i.e. DOING	VDG	62	0.4
past participle of the verb "BE", i.e. BEEN	VBN	45	0.3
-ing form of the verb "HAVE", i.e. HAVING	VHG	42	0.3
past participle of the verb "DO", i.e. DONE	VDN	20	0.1
past tense form of the verb "HAVE", i.e. HAD, 'D	VHD	20	0.1
past form of the verb "DO", i.e. DID	VDD	10	0.1
past participle of the verb "HAVE", i.e. HAD	VHN	3	0.0
Total		14409	

Table 4 shows the distribution of prepositions, adjectives and conjunctions in the learner corpus ordered respectively according to their frequencies. Among the prepositions 76% is miscellaneous prepositions and nearly 24% is the preposition 'of'. Among the adjectives, 63% is unmarked adjectives while 35% is comparative and only nearly 2% is superlative adjectives. With regard to conjunctions, nearly 70 percent of the conjunctions used by learners are subordinating conjunctions, 25% is the conjunction 'that' and only 5% is coordinating conjunctions. This distribution indicates that the learners prefer to use complex sentences rather than simple sentences in their essays, perhaps because of the fact that they were practicing with writing academic essays for the tasks that led to the creation of the corpus.

Table 4. Distribution of prepositions, adjectives and conjunctions in the learner corpus

Types	Codes	Frequency	%
preposition (except for OF) (e.g. FOR, ABOVE, TO)	PRP	6135	76.1
the preposition OF	PRF	1927	23.9
Total		8062	
adjective (unmarked) (e.g. GOOD, OLD)	AJO	4990	63.2
comparative adjective (e.g. BETTER, OLDER)	AJC	139	35.0
superlative adjective (e.g. BEST, OLDEST)	AJS	105	1.8
Total		7891	
subordinating conjunction (e.g. ALTHOUGH, WHEN)	CJS	1380	69.1
the conjunction THAT	CJT	513	25.7
coordinating conjunction (e.g. AND, OR)	CJC	2762	5.3
Total		1998	

4.1 Classification of Errors

The essays in the error corpus were hand-tagged by using the error classification system which is a combination of error categories from GRADES (Bowden & Fox, 2002) and Error tagging Manual, version 1.1. (Dagneaux, Denness, Granger, & Meunier, 1996) and the additional error categories were included according to the task requirements and the specific requirements of the corpora felt by the researcher. The hand tagging of the errors in the learner corpus was done by two trained independent raters. After the two independent raters completed their tagging procedure, an interrater reliability coefficient was calculated between the ratings using the Cohen's Kappa interrater reliability measure. The resulting interrater reliability was found to be 0.86 which is quite high indicating that the error categories were effective in classifying the error types. After the hand tagging was completed, the produced error corpus was analyzed using a concordancing program, namely AntConc 3.2.4. (Anthony, 2014).

Table 5. Major error categories used in the study

Error Category	Number of related error types
Verb related errors	15
Clausal errors	8
APA style related errors	7
Noun related errors	6
Pronoun related errors	3
Preposition related errors	3
Adjective related errors	3
Word choice	3
Adverb related errors	1
Mechanics	1
Phrase choice	1
	51

4.2 Comparison of Error Types

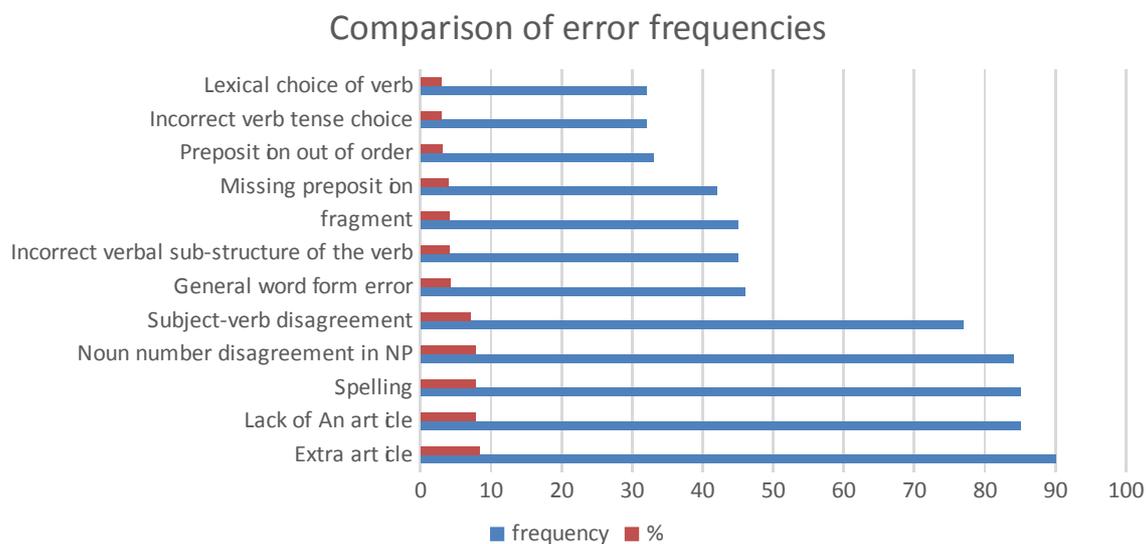


Figure 3. Comparison of error frequencies

As a result of the tagging of errors, it was possible to see the most frequent individual error types for Turkish EFL learners when writing academic essays. In terms of frequency, the top error types when considered individually are shown in Figure 3. The two most frequent error types are both related to the use of a determiner before a noun: using an extra article or omitting an article where necessary. This error type can be considered as an interference error since Turkish language does not have an article system unlike the English language. These errors account for 8.3% (extra article use) and 7.8% (lack of an article) of all the errors. Following article errors, the third most frequent error type among all errors is spelling errors which account for 7.8% of all the errors. Although spelling errors are not very serious errors, in academic writing they can be seen as important since they affect the credibility of a writer. The frequency of spelling errors shows the lack of word processing skills or lack of proofreading which are important skills for a writer to develop. Among the most frequent errors, the fourth and fifth most frequent errors are noun number disagreement in NP (7.8%) and subject verb agreement errors (7.1%). These errors also result from first language influence and are therefore interference errors. Within the most frequent 12 error types listed in Figure 2, there are four errors which relate to the use of verbs which indicates that Turkish EFL learners have difficulty with the use of verbs in writing mostly. For example, subject-verb agreement (7.1%), incorrect verbal sub-structure of the verb (4.2%), incorrect verb tense choice (3%) and lexical choice of verb (3%). Other problematic areas are word forms (4.2%), fragments (4.2%), missing prepositions (3.9%) and prepositions used wrongly (3%).

In the next section, the errors are discussed in more detail under categories with example sentences from the corpus.

4.3 Verb Related Errors

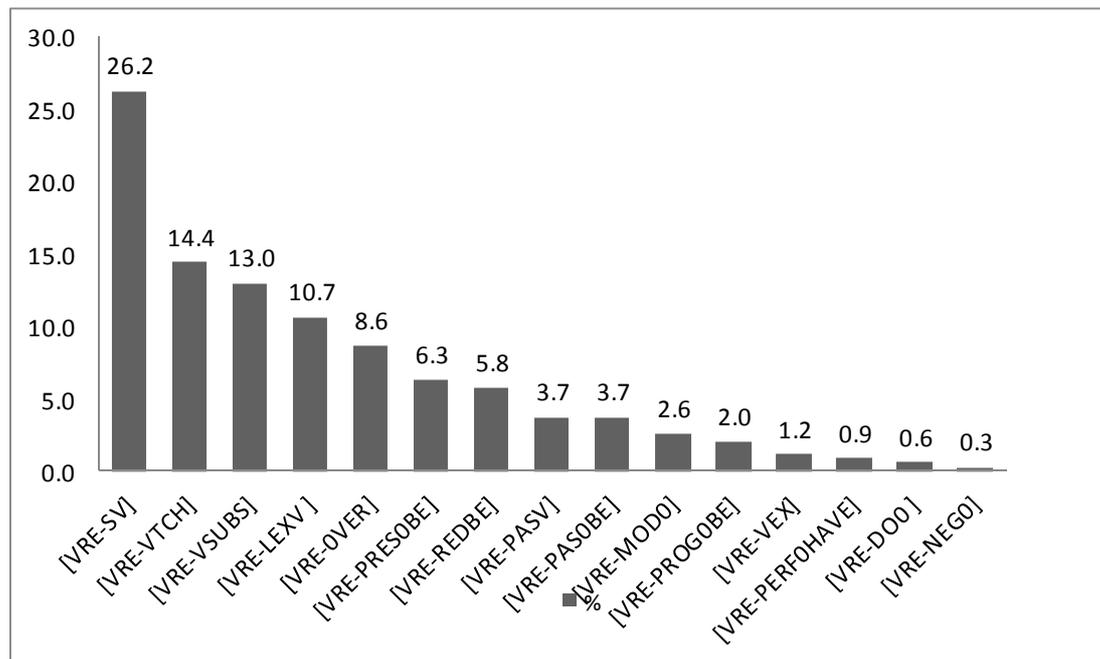


Figure 4. Distribution of verb-related errors

The most frequently observed verb related errors in the corpus were related to subject verb agreement (26.2%), verb-tense choice (14.4%), verbal sub-structure of the verb (13%) and lexical choice of verb. An examination of the erroneous sentences (i.e. 1a, 1b) shows that students mainly have difficulty with the subject verb agreement rules in simple present tense; most commonly failing to use 3rd person singular grammatical morpheme. They also have difficulty with the agreement of auxiliary verbs have and do in the simple present tense. These errors can be classified as interlingual transfer errors since Turkish language does not have inflection in the third person singular.

Subject verb agreement

(1a) File 14.argu.ess.002: In one interview, they asked the people about piracy and those who [VRE-SV] downloads from internet...

(1b) File 14.argu.ess.031: ... if this student [VRE-SV] do not control time for examinations, he or she will have problems ...

(1c) File 14.argu.ess.032: Every historian [VRE-SV] have different ideas about it.

Verb tense choice

(1d) File 14.argu.ess.011: For example; when a producer makes a film he [VRE-VTCH] sent the film to the cinemas to show the film.

Lexical Choice of Verb

(2a) File 14.argu.ess.013: Firstly, piracy is [VRE-LEXV] being more common in society because it provides getting information or collecting data more quickly and that's why people tend to use this.

(2b) File 14.argu.ess.040: There are some reasons why the stress occurs when they

[VRE-LEXV] present a presentation.

4.4 Noun Related Errors

Most frequent noun-related error is redundant use of an article with a noun (32%) followed by missing article (28%), noun number mistakes (27.6%), lexical choice of noun (10.5%). Both article errors; redundant use of article and missing article could be considered as interference errors since Turkish language does not have an

article system like that of the English language.

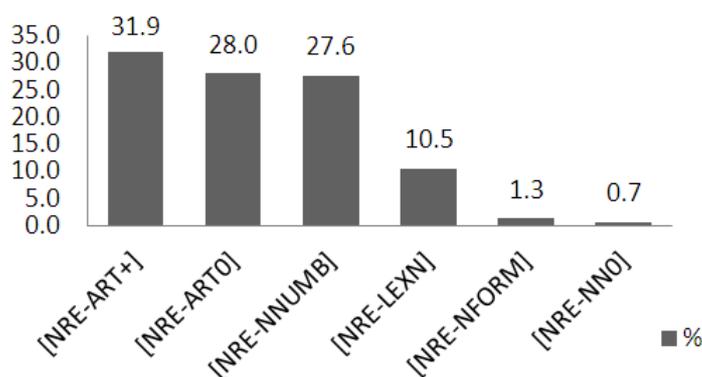


Figure 5. Distribution of noun-related errors

Redundant use of article

(3a) File 14.defi.ess.015: If there is [NRE-ART+] a something that it is sorrowful, both of you will be upset ...

Missing article

(4a) File edu.ays.alt: Therefore, education is [NRE-ART0] basic necessity for [NRE-ART0] development a country but it can be [WF] total opposite.

Noun number

(5a) File edu.su.he: From past to now, a lot of [NRE-NNUMB] economist, social and political [NRE-NNUMB] scientist put forward that education plays a huge and important role in development as a country.

Lexical choice of noun

(6a) File 14.argu.ess.011: Because of the internet piracy nobody watches the [NRE-LEXN] clips on television ...

4.5 Adjective Related Errors

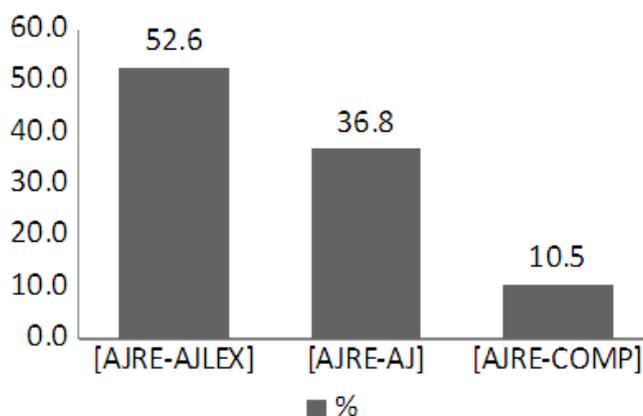


Figure 6. Distribution of adjective-related errors

Nearly half of the adjective related errors observed in the corpus are lexical choice errors (52.6%) which could be classified as conceptual errors. Example (6a) displays this error type. The students fail to match concepts to vocabulary items if they cannot be directly translated from L1 to L2. The error category “adjective out of order” which accounts for 36.8% of all the adjective errors defines adjectives which are made up by students through false generalizations such as in example (7a) below; ‘institutive’ or ‘systematical’. The errors relating the comparative form which accounts for 10.5% or the adjective errors, such as in example (8a) could be considered

as intralingual errors due to failures in the learners' internalization of syntactic rules of English.

Lexical choice of adjective

(6a) File 14.proc.ess.039: To sum up, to control your stress is easier than you think. The only thing you have to do is to [VRE-VSUBS] relief your mind, take a [AJRE-AJLEX] large breath and find a solution to struggle with the stress in your own way.

Adjective out of order

(7a) File edu. bed.oz.: Sumarians were the first people setting institutive [AJRE-AJ] and systematical [AJRE-AJ] education customs as first people who invented the writing among the other human societies.

Incorrect use of the comparative/superlative form

(8a) File edu.ays.alt. : Thanks to education, countries have more better [AJRE-COMP] conditions for their society

4.6 Style (APA) Errors

Since the writing task required students to follow the APA style, errors relating the use of APA style were also included in the analysis. These errors included both in-text citation errors and errors which affect academic writing style such as wordiness.

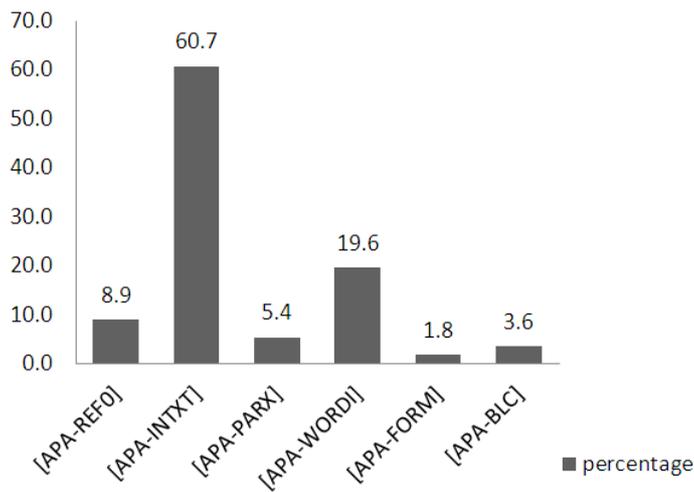


Figure 7. Distribution of APA-style related

A majority of the APA-style related errors were in-text citation errors (60.7%) such as the one in example (9a). In this example, although the writer has acknowledged the source material, he/she has failed to cite the source properly; that is using last name and parenthetical citation. The frequency of these errors is an indication that students need more training in in-text citation rules. The second most frequent APA-style error is wordiness (19.6%); that is repeating words which are near in meaning. Students might be making this mistake for the sake of making their sentences longer and seemingly more complex. Example (9c) shows this kind of error. In this example both “contemporary” and “present” are used in the sentence one after the other although they are synonymous. Being able to paraphrase source material is an important component of academic writing. Therefore, paraphrases were also examined, and weak paraphrases (5.4%) such as in example (9b) were included in the analysis. However, the frequency of these types of error is relatively low.

(9a) File 14.defi.ess.017: According to [APA-INTXT] Dr. Suzan Akkaya, a psychologist in Marmara University, most of the lung cancer sufferers there are a big chance to live in today's world...

(9b) File 14.proc.ess.033: Mark Zuckerberg, one of the co-founders of Facebook, said that [APA-PARX]“their company's mission was make the world more open and connected”. Some say privacy policy is not enough to protect people from some crimes, but if you know how to rearrange your privacy, there will be no crime for you.

(9c) File edu.alp.mut: Furthermore, when we come to our [APA-WORDI] contemporary present day, ...

4.7 Clausal Errors

Clausal errors included in the analysis were: fragment [CLE-FRAG], Run-on sentence [CLE-RUNON], incorrect word order in embedded clause [CLE-WOX], choice of clausal connector [CLE-CLACX], redundant clausal connector [CLE-CLAC+], lack of noun clause marker [CLE-CM0], lack of adjective clause marker [CLE-AJCX], and lack of a subject in a clause [CLE-SUB0]. Figure 8 shows the distribution of clausal errors in the corpus.

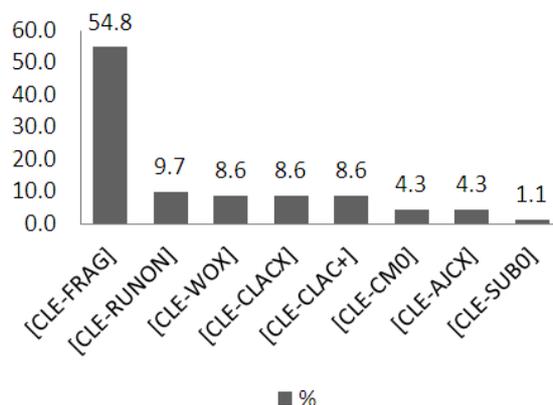


Figure 8. Distribution of clausal errors

An examination of clausal errors revealed that approximately half of all the clausal errors are fragments with 54.8%. In example (10a), the fragment is caused by separating the subordinate clause from the main clause and starting the subordinate clause which a subordinator with a capital letter. The second most frequent error type in this category is run-on sentence (9.7%). Example (10b) shows a run-on sentence error.

(10a) File 14.argu.ess.012: economic condition of country. [CLE-FRAG] Because they know that if the economy of the country is not high, their income will decline after a while.

(10b) File 14.argu.ess.012: If something or someone is bothering you, communicate your concerns in an open and respectful way [CLE-RUNON] manage your time better.

4.8 Examples of Other Kinds of Errors

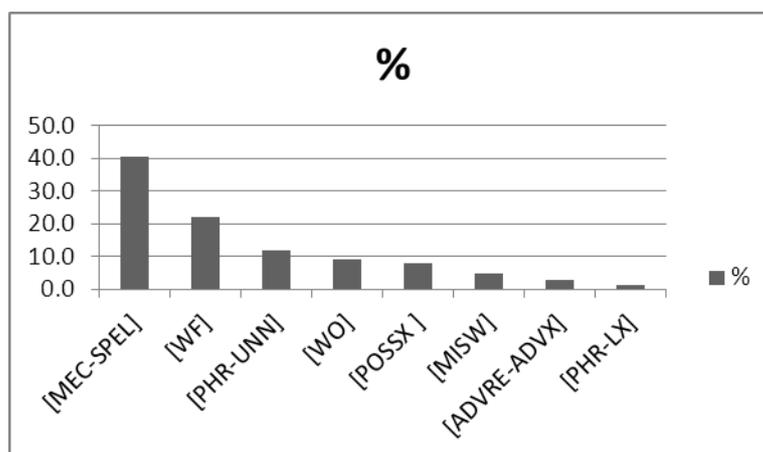


Figure 9. Distribution of other types of errors

The category of other forms of error was treated as a different category which included word form, unnecessary word or phrase, general word order and possessive form errors. In this category spelling errors accounted for 40.4% of all the errors. 22.1% of the errors were related to the use of word forms like the error shown in example

(11a) where an adjective is used instead of an adverb and thus the wrong word form is selected. 11.7% of errors were caused by the use of an unnecessary word or phrase like the error displayed in example (11b). In this example from the corpus the indefinite pronoun “it” is unnecessary since the sentence already has a subject “stress” and additionally in the relative clause the relative clause marker “that” modifies the subject. Word order errors such as the one in example (11c) accounted for 8.9% of the errors in this category. In this example, the place of the adverb “really” should be before the verb but it is wrongly placed before the subject “I”. This could be acceptable in spoken language but since this example comes from written production it was not accepted as correct. In the last example,

Word form

(11a) File 14.argu.ess.001: They make some films and songs and their expectation rises both economically and [WF] social (socially) ; however, ...

Unnecessary word or phrase

(11b) File 14.defi.ess.007: Stress is a situation that, [PHR-UNN] it happens when people feel under pressure [PHR-UNN] themselves.

Word order

(11c) File 14.defi.ess.016: When I was confronted with this problem, [WO] really I was shocked.

Possessive forms

(11d) File 14.argu.ess.002: One of [POSSX] thoughts of them is while we can access from

5. Conclusion

As a result of the analysis of errors in the error tagged corpus yielded the following results. The three areas that are most problematic for the students are verbs, nouns and prepositions. Verb use related errors account for 26.3% of all language errors. Among verb related errors, the omission or redundant use of the 3rd person singular grammatical morpheme by Turkish EFL students has not been widely researched; however, in studies by Ertekin (2006) and Ülgü et al. (2013), these errors have been classified by interlanguage transfer errors. This is due to the fact that Turkish language does not have inflection in the third person singular. Similar problems with the use of this morpheme have been reported in Chinese EFL learners as well (Hsiesh, 2009).

Noun use related errors account for 23% of all language errors. These errors are redundant use of articles, omission of articles, noun number errors and lexical choice of nouns. Most frequent errors in this category are noun modification errors relating the use of articles. These errors can be classified by interlingual transfer errors since Turkish language does not have articles which modify nouns. As for the noun number errors, as Erdoğan (2005) also highlights, Turkish students tend to omit the plural suffix at the end of the word as Turkish does not require its use in adjectival phrases. Errors in lexical choice are also due to transfer of some lexical items into the target language.

Preposition use related errors accounts for 15.7% of all the errors. The errors in the use of prepositions have been regarded by Turkish researchers as interlingual transfer errors as highlighted by Karatay (2011) in a comprehensive analysis of syntactic errors in the compositions of Turkish EFL students written for a proficiency exam. The least frequent errors are made in the categories of missing word (0.8), adverbs (0.5) and lexical choice of phrases (0.2). There is a relationship between the frequency of use of language items and errors performed in using them as shown by the frequencies obtained from the tagged version of the corpus since nouns, verbs and prepositions were found to be the most frequently used items in the learner corpus.

Style related errors which concern the use of the APA style account for 4% of all errors. Among these errors, the most frequent ones are in-text citation format errors (60%), wordiness (19.6%) and lack of reference for borrowed information of concept (9%). The APA style rules are concepts that are novel for the EFL learners since they have heard about these rules only in the context of the academic writing course. Therefore, rules related to the use of APA style rules could be classified as developmental errors as defined by Richards (1974) and more specifically incomplete of application of rules.

6. Practical Implications

The corpus analysis carried out in this study and the method carried out for the analysis carry an important potential for language professionals since it provides valuable data for the planning of teaching activities and design of materials which cater for the needs of learners. CEA studies on different EFL groups will help future

studies on error tagging by defining specific error types encountered in learner language of different native language backgrounds. These studies will contribute to more effective computer based analysis of non-native language production. On the pedagogic side, CEA studies will guide teaching practices and the development of language materials which focus specifically on error types which are more probable to be performed by certain learner groups. In this way, teachers may focus on potentially problematic areas for specific learner groups with different language backgrounds. Language materials may also better cater for the language learning needs of specific learner groups.

Teachers of writing have to take on multiple responsibilities in that they have to teach many skills including how to organize an academic paper, how to carry out research on a topic, how to follow an academic style when writing but they are not necessarily teaching the target language. Regarding the language problems emerging in students' writing, the teachers should train the student writers to take on more responsibility for their own work and to act as their own critiques to prevent errors such as subject verb agreement errors or article use errors. In terms of L1 induced errors, as Bestgen et al. (2012) put it as a result of error analyses, detected L1 salient error types can help teachers zoom in on learner-corpus-attested areas of difficulty and especially advanced learners could try and erase the more visible signs of their nonnativeness from their writing.

References

- Adjedi, A. A. (2015). Analysis of subordination errors in students' writings: A study of selected teacher training colleges in Ghana. *Journal of Education and Practice*, 6(8), 62-77.
- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-195. <https://doi.org/10.1093/applin/22.2.173>
- Anthony, L. (2014). AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from <http://www.laurenceanthony.net>
- Bestgen, Y, Granger, S., & Thewissen, J. (2012). Error patterns and automatic L1 identification. In S. Jarvis, & S. A. Crossley (Eds.), *Approaching Language Transfer Through Text Classification: Explorations in the Detection-based Approach* (pp. 127-153). Bristol: Multilingual Matters.
- Bowden, M. I., & Fox, R. K. (2002). *A diagnostic approach to the detection of syntactic errors in English for non-native speakers*. The University of Texas–Pan American Department of Computer Science Technical Report.
- Chuang, F-Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of Chinese student writing. *Corpora*, 1, 251-271. <https://doi.org/10.3366/cor.2006.1.2.251>
- Crossley, S. A., & McNamara, D. S. (2012). Detecting the first language of second language writers using automatic indices of cohesion, lexical sophistication, syntactic complexity and conceptual knowledge. In S. Jarvis, & S. A. Crossley (Eds.), *Approaching language transfer through text classification: Explorations in the detection-based approach* (pp. 127-153). Bristol: Multilingual Matters.
- Dagneaux, E., Denness, S., Granger, S., & Meunier, F. (1996). *Error tagging manual version 1.1*. Louvain-la Neuve: Centre for English Corpus Linguistics, Université Catholique de Louvain.
- Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided error analysis. *System*, 26, 164-174. [https://doi.org/10.1016/S0346-251X\(98\)00001-3](https://doi.org/10.1016/S0346-251X(98)00001-3)
- Díez-Bedmar, M. B. & Papp, S. (2008). The use of the English article system by Chinese and Spanish learners. In G. S. Gilquin, S. Papp, & M. B. Díez-Bedmar (Eds.) *Linking Up: Contrastive and Learner Corpus Research*. (pp. 147-175). Amsterdam: Rodopi. https://doi.org/10.1163/9789401206204_007
- Díez-Bedmar, M. B. (2011). Spanish pre-university students' use of English: CEA results from the University Entrance Examination, *International Journal of English Studies*, 11(2), 141-158.
- Doolan, S. M., & Miller, D. (2012). Generation 1.5 written error patterns: A comparative study. *Journal of Second Language Writing*, 21(1), 1-22. <https://doi.org/10.1016/j.jslw.2011.09.001>
- Erdoğan, V. (2005). Contribution of error analysis to foreign language teaching, *Mersin University Journal of the Faculty of Education*, 1(1), 261-270.
- Erkaya, O. R. (2012). Vocabulary and L1 interference–error analysis of Turkish students. *Literacy Issues in Higher Education*, 36(2), 1-11.
- Ertekin, D. (2006). *Acquisition of L2 English inflectional system by Turkish native speakers: Rule-governed or rote-learned*. Unpublished Dissertation, Çukurova University, Adana, Turkey.

- Ferris, D. R. (2004). The “grammar correction” debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime...?). *Journal of second language writing*, 13(1), 49-62. <https://doi.org/10.1016/j.jslw.2004.04.005>
- Flowerdew, J. (2010). Use of signaling nouns across L1 and L2 writer corpora. *International Journal of Corpus Linguistics*, 15, 36-55. <https://doi.org/10.1075/ijcl.15.1.02flo>
- Garside, R. (1987). The CLAWS Word-tagging System. In R. Garside, G. Leech, & G. Sampson (Eds.), *The Computational Analysis of English: A Corpus-based Approach*. London: Longman.
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, 32(3), 301-319. <https://doi.org/10.1016/j.system.2004.04.001>
- Granger, S., & Meunier, F. (1994). Towards a grammar checker for learners of English. In U. Fries, & G. Tottie, (Eds.). *Creating and Using English Language Corpora*. (pp. 79-91). Amsterdam: Rodopi.
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non - native EFL speakers of English. *World Englishes*, 15(1), 17-27. <https://doi.org/10.1111/j.1467-971X.1996.tb00089.x>
- Han, N., Chodorow, M., & Leacock. (2006). Detecting errors in English article usage by non-native speakers, *Natural Language Engineering*, 12(2), 115-129. <https://doi.org/10.1017/S1351324906004190>
- Hawkins, J., & Buttery, P. (2010). Criterial features in learner corpora: Theory and illustrations. *English Profile Journal*, 1, 1-23. <https://doi.org/10.1017/S2041536210000036>
- Hsieh, F. T. (2009). *The acquisition of English tense/agreement morphology and copula be by L1-Chinese-speaking learners*. Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching, Lancaster: Lancaster University, pp. 46-59.
- Huaqing, H. (2016). A Computer-Aided Analysis of Word Form Errors in College English Writing –A Corpus-based Study. *International Journal of Engineering and Techniques: iJET*, 11(3), 1-7.
- Karatay, H. (2011). 4+ 1 planlı yazma ve değerlendirme modelinin öğretmen adaylarının yazılı anlatım tutumlarını ve yazma becerilerini geliştirmeye etkisi. *Turkish Studies*, 6(3), 1029-1047.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*, New York: Addison Wesley Longman Inc.
- Kırkgöz, Y. (2010). An analysis of written errors of Turkish adult learners of English. *Procedia-Social and Behavioral Sciences*, 2(2), 4352-4358. <https://doi.org/10.1016/j.sbspro.2010.03.692>
- Mammeri, S. (2015). A Morphosyntactic Study of EFL Students’ Written Compositions: A Corpus Based Analysis, *Arab World English Journal*, 112-126.
- Milton, J., (1994). A corpus-based online grammar and writing tool for EFL learners: a report on work in progress. In A. Wilson, & T. McEnery (Eds.). *Corpora in Language Education and Research*. (pp. 65-77). Unit for Computer Research on the English Language, Lancaster University.
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp.125-156). Oxford: John Benjamins Publishing Company. <https://doi.org/10.1075/scl.17.08nes>
- Nezami, A., & Najafi, M. S. (2012). Common error types of Iranian learners of English. *English Language Teaching*, 5(3), 160-170. <https://doi.org/10.5539/elt.v5n3p160>
- Özhan, D. (2012). A comparative analysis on the use of but, however and although in the university students’ argumentative essays: A corpus-based study on Turkish learners of English and American native speakers. *Unpublished Doctoral Dissertation*. METU, Ankara, Turkey.
- Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, 18, 391-417. <https://doi.org/10.1075/ijcl.18.3.06paq>
- Punga, L., & Parlog, H. (2015). ‘He is a criminal in series’: a foray into errors by Romanian learners of English. *Professional Communication and Translation Studies*, 8, 161-156
- Richards, J. (1974). ‘A Non-Contrastive Approach to Error Analysis’. In J. Richards (Ed.). *Error analysis: Perspectives on Second Language Acquisition* (pp. 172-188). Essex: Longman.
- Thewissen, J. (2013). Capturing L2 Accuracy Developmental Patterns: Insights From an Error - Tagged EFL Learner Corpus. *The Modern Language Journal*, 9777-101. <https://doi.org/10.1111/j.1540-4781.2012.01422.x>

- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of second language Writing*, 16(4), 255-272. <https://doi.org/10.1016/j.jslw.2007.06.003>
- Ülğü, S., Nisanci, S., & Ünal, S. (2013). Acquisition of Third Person Singular-s: A Primary State School (Sixth Graders) Case in Turkey. *Procedia-Social and Behavioral Sciences*, 70, 1347-1353. <https://doi.org/10.1016/j.sbspro.2013.01.196>
- Xia, L. (2015). An Error Analysis of the Word Class: A Case Study of Chinese College Students. *International Journal of Engineering and Techniques*, 10(3), 41-45.

Appendix A

UCREL CLAWS5 Tagset

- AJ0 adjective (unmarked) (e.g. GOOD, OLD)
- AJC comparative adjective (e.g. BETTER, OLDER)
- AJS superlative adjective (e.g. BEST, OLDEST)
- AT0 article (e.g. THE, A, AN)
- AV0 adverb (unmarked) (e.g. OFTEN, WELL, LONGER, FURTHEST)
- AVP adverb particle (e.g. UP, OFF, OUT)
- AVQ wh-adverb (e.g. WHEN, HOW, WHY)
- CJC coordinating conjunction (e.g. AND, OR)
- CJS subordinating conjunction (e.g. ALTHOUGH, WHEN)
- CJT the conjunction THAT
- CRD cardinal numeral (e.g. 3, FIFTY-FIVE, 6609) (excl ONE)
- DPS possessive determiner form (e.g. YOUR, THEIR)
- DT0 general determiner (e.g. THESE, SOME)
- DTQ wh-determiner (e.g. WHOSE, WHICH)
- EX0 existential THERE
- ITJ interjection or other isolate (e.g. OH, YES, MHM)
- NN0 noun (neutral for number) (e.g. AIRCRAFT, DATA)
- NN1 singular noun (e.g. PENCIL, GOOSE)
- NN2 plural noun (e.g. PENCILS, GEESE)
- NP0 proper noun (e.g. LONDON, MICHAEL, MARS)
- NULL the null tag (for items not to be tagged)
- ORD ordinal (e.g. SIXTH, 77TH, LAST)
- PNI indefinite pronoun (e.g. NONE, EVERYTHING)
- PNP personal pronoun (e.g. YOU, THEM, OURS)
- PNQ wh-pronoun (e.g. WHO, WHOEVER)
- PNX reflexive pronoun (e.g. ITSELF, OURSELVES)
- POS the possessive (or genitive morpheme) 'S or '
- PRF the preposition OF
- PRP preposition (except for OF) (e.g. FOR, ABOVE, TO)
- PUL punctuation - left bracket (i.e. (or [)

PUN	punctuation - general mark (i.e. . ! , ; - ? ...)
PUQ	punctuation - quotation mark (i.e. ` ' “ ”)
PUR	punctuation - right bracket (i.e.) or])
TO0	infinitive marker TO
UNC	“unclassified” items which are not words of the English lexicon
VBB	the “base forms” of the verb “BE” (except the infinitive), i.e. AM, ARE
VBD	past form of the verb “BE”, i.e. WAS, WERE
VBG	-ing form of the verb “BE”, i.e. BEING
VBI	infinitive of the verb “BE”
VBN	past participle of the verb “BE”, i.e. BEEN
VBZ	-s form of the verb “BE”, i.e. IS, 'S
VDB	base form of the verb “DO” (except the infinitive), i.e.
VDD	past form of the verb “DO”, i.e. DID
VDG	-ing form of the verb “DO”, i.e. DOING
VDI	infinitive of the verb “DO”
VDN	past participle of the verb “DO”, i.e. DONE
VDZ	-s form of the verb “DO”, i.e. DOES
VHB	base form of the verb “HAVE” (except the infinitive), i.e. HAVE
VHD	past tense form of the verb “HAVE”, i.e. HAD, 'D
VHG	-ing form of the verb “HAVE”, i.e. HAVING
VHI	infinitive of the verb “HAVE”
VHN	past participle of the verb “HAVE”, i.e. HAD
VHZ	-s form of the verb “HAVE”, i.e. HAS, 'S
VM0	modal auxiliary verb (e.g. CAN, COULD, WILL, 'LL)
VVB	base form of lexical verb (except the infinitive)(e.g. TAKE, LIVE)
VVD	past tense form of lexical verb (e.g. TOOK, LIVED)
VVG	-ing form of lexical verb (e.g. TAKING, LIVING)
VVI	infinitive of lexical verb
VVN	past participle form of lex. verb (e.g. TAKEN, LIVED)
VVZ	-s form of lexical verb (e.g. TAKES, LIVES)
XX0	the negative NOT or N'T
ZZ0	alphabetical symbol (e.g. A, B, c, d)

Appendix B

Error Categories Used in the Study

Type of error

Subject-verb disagreement
Incorrect verb tense choice
Incorrect verbal sub-structure of the verb

Error code

[VRE-SV]
[VRE-VTCH]
[VRE-VSUBS]

Lexical choice of verb	[VRE-LEXV]
Lack of the main verb	[VRE-0VER]
Lack of the auxiliary verb 'be' for the present tense	[VRE-PRES0BE]
Redundant use of auxiliary verb 'be'	[VRE-REDBE]
Incorrect use of passive voice	[VRE-PASV]
Lack of the auxiliary verb 'be' for the passive voice	[VRE-PAS0BE]
Lack of modal auxiliary	[VRE-MOD0]
Lack of the auxiliary verb 'be' for the progressive tense	[VRE-PROG0BE]
Verbal elements out of order	[VRE-VEX]
Lack of the auxiliary verb 'have' for the perfect tense	[VRE-PERF0HAVE]
Lack of the auxiliary verb 'do' for the present tense	[VRE-DO0]
Lack of negation	[VRE-NEG0]
Extra article	[NRE-DET+]
Lack of An article	[NRE-DET0]
Noun number disagreement in NP	[NRE-NNUMB]
Lexical Choice of noun	[NRE-LEXN]
Noun form error	[NRE-NFORM]
Missing noun	[NRE-NN0]
Lack of reference	[APA-REF0]
In-text citation format	[APA-INTXT]
Weak paraphrase or translation	[APA-PARX]
Reference page format errors	[APA-REFPG]
wordiness	[APA-WORDI]
Formality	[APA-FORM]
Block quotation required but borrowed info. presented like a paraphrase	[APA-BLC]
Pronoun choice error	[PRO-PRX]
Pronoun missing	[PRO-PRon0]
Redundant pronoun	[PRO-Pron+]
Preposition out of order	[PRE-PRP]
Redundant preposition	[PRE-PRE+]
Missing preposition	[PRE-PRP0]
Lexical Choice of adjective	[AJRE-AJLEX]
Pronoun missing	[PRO-PRon0]
Redundant pronoun	[PRO-Pron+]
Preposition out of order	[PRE-PRP]
Redundant preposition	[PRE-PRE+]
Missing preposition	[PRE-PRP0]
Lexical Choice of adjective	[AJRE-AJLEX]
Adjective out of order	[AJRE-AJ]
incorrect use of the comparative form of adjective	[AJRE-COMP]
Incorrect use of adverb	[ADVRE-ADVX]
fragment	[CLE-FRAG]

Run-on sentence	[CLE-RUNON]
Incorrect word order in embedded clause	[CLE-WOX]
Choice of clausal connector	[CLE-CLACX]
Redundant clausal connector	[CLE-CLAC+]
Lack of noun clause marker	[CLE-CM0]
Lack of adjective clause marker	[CLE-AJCX]
Lack of a subject in a clause	[CLE-SUB0]
Spelling	[MEC-SPEL]
General word form error	[WF]
Unnecessary phrase or word	[PHR-UNN]
Misplaced word	[WO]
Possessive form error	[POSSX]

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).